ESC-TR-93-244

Technical Report
982

Use of the Weiss-Weinstein Bound to Compare the Direction-Finding Performance of Sparse Arrays

D.F. DeLong

16 August 1993

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS

Prepared for the Department of Defense under Air Force Contract F19628-90-C-0002.

This report is based on studies performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. The work was sponsored by the Department of Defense under Air Force Contract F19628-90-C-0002.

The ESC Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Gary Tutungian
Administrative Contracting Officer
Directorate of Contracted Support Management

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

USE OF THE WEISS-WEINSTEIN BOUND TO COMPARE THE DIRECTION-FINDING PERFORMANCE OF SPARSE ARRAYS

D.F. DeLONG
Group 44

TECHNICAL REPORT 982

16 AUGUST 1993

LEXINGTON MASSACHUSETTS

ABSTRACT

Weiss and Weinstein have recently derived a new lower bound on random parameter estimation error. This report makes use of this bound to develop a criterion for comparing two sparse array configurations. The arrays are first adjusted in size so that their direction-finding (DF) accuracies at high element signal-to-noise ratio (ESNR) are the same when averaged over all possible directions of arrival in the hemisphere. Then their threshold ESNRs are calculated using the Weiss-Weinstein bound. The array having the lower threshold ESNR is judged the better array. As an example of the application of the bound, three 7-element arrays, a circle, a spiral, and an L-shaped array, are compared.

The bound is also extended to the case of switched subarrays, where data are collected from several subarrays sequentially in time and processed as a unit. The purpose is to achieve DF accuracy approaching that of the full array while reducing the number of signal processing channels. An example, using five receivers on the 7-element spiral array, shows that this is indeed possible.

ACKNOWLEDGMENTS

The author wishes to thank Irv Stiglitz for raising the issues out of which this work grew. Thanks are due equally to Larry Horowitz and Jay Sklar for the support and guidance they provided during its execution and also for their painstaking reviews of the manuscript, which greatly improved its exposition. The mathematical guidance provided by Keith Forsythe is also gratefully acknowledged.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

LIST OF ILLUSTRATIONS

1. INTRODUCTION

The purpose of this report is to develop a method for comparing two sparse planar array configurations. The criterion for comparison is based on narrowband single-signal direction-finding (DF) accuracy in the presence of spatially white noise (i.e., uncorrelated from receiver to receiver). The directions of arrival are assumed to be confined to a hemisphere, and DF error is averaged over all possible directions of arrival in the hemisphere.

According to sampling theory, the aperture illumination must be sampled about every half-wavelength in both spatial dimensions in order to preserve all the available information about the directions of the signals. Each sample requires an antenna element and a receiver channel. Unfortunately, system cost increases very rapidly with the number of receiver channels used in the array.

The performance of an array is characterized by its array pattern, which is the response of the array, phased to point in a particular direction, to a signal arriving from another direction. When the array satisfies the half-wavelength sampling criterion, the array pattern has a well-defined main beam and uniformly low sidelobes. As elements are removed from the array, the sidelobes grow in size and number. A high sidelobe means that there are two directions having very similar array response vectors. In the presence of noise and calibration uncertainties, confusion between the mainlobe and the sidelobes can result. If the number of signals is known, or can be reliably estimated, this confusion manifests itself as occasional large DF errors; if not, it causes false alarms as well.

The value of element signal-to-noise ratio (ESNR) at which these large errors begin to occur with significant probability is called the threshold ESNR of the array. If two arrays have equal accuracy at high ESNR, the array having the lower threshold ESNR is deemed the better array.

The Weiss-Weinstein (WW) bound is a lower bound on the mean squared estimation error for random parameters with known *a priori* distributions. As will be shown, it predicts remarkably well the threshold ESNR of the maximum likelihood direction estimator. This quantity would otherwise have to be determined by simulation.

Equations for the bound are developed for the planar array problem, and various issues that arise in its calculation are addressed. A scalar measure of accuracy is proposed for two-dimensional DF problems, and the *Cramer-Rao (CR) factor* of an array is defined. Arrays having the same CR factor have the same DF accuracy bound at high ESNR. As an example of the application of the bound, three 7-element arrays, a circle, a spiral, and an L-shaped array, are compared.

The bound is extended to the case of *switched subarrays*, where data is collected from several subarrays sequentially in time and processed as a unit. The purpose is to achieve DF accuracy approaching that of the full array while reducing the number of signal processing channels. An example, using five receivers on the 7-element circular array, shows that this is indeed possible.

The main advantage of adaptive arrays is their ability to adapt their array patterns to the signal environment, amplifying the signals of interest while simultaneously reducing other unwanted signals. Nevertheless, as a first step toward a rationale for adaptive array comparison, arrays will be compared on the basis of their single-signal DF accuracy, assuming ideal (i.e., omnidirectional) antenna element patterns. If this approach provides some insight into the problem, then perhaps it can be extended to more complicated scenarios.

The following notational conventions are used. Bold lowercase letters represent (column) vectors, while bold uppercase letters represent matrices. No particular distinction is made between real and complex quantities; the context should make it clear. The transpose of a matrix $\mathbf{A}$ is $\mathbf{A}^T$; its complex conjugate is $\mathbf{A}^*$; its conjugate, or Hermitian, transpose is $\mathbf{A}^H$. The determinant of $\mathbf{A}$ is denoted by $\det \mathbf{A}$. $\mathbf{I}$ denotes the identity matrix; $\mathbf{1}$ denotes a vector having all elements equal to 1. The notation $[\mathbf{x}]$ denotes the diagonal matrix, the nonzero elements of which are those of the vector $\mathbf{x}$. The inequality $\mathbf{A} \geq \mathbf{B}$ means that the matrix $\mathbf{A} - \mathbf{B}$ is nonnegative definite. All matrix inverses (including scalars) that appear are assumed to exist, unless otherwise noted.

2. THE PARAMETER ESTIMATION PROBLEM

DF is an example of a parameter estimation problem. An excellent discussion of such problems is available in Van Trees [1]. The following brief description is given primarily to introduce some terminology and notation.

Observations of a (scalar- or vector-valued) random process are made in order to determine which of a set of possible probability density functions (pdfs) generated them. If the set of possible pdfs is finite (usually small), the problem is called a hypothesis testing problem; otherwise, it is called a parameter estimation problem. *Parameters* are quantities that may be thought of as indexing the set of possible pdfs, i.e., knowing the values of the parameters is equivalent to knowing which pdf is true. The vector, the elements of which are the parameter values, is called the *parameter vector*.

To best model the physical situation at hand, the elements of the parameter vector may be treated either as unknown constants or as random variables with a joint pdf of their own. When they are treated as unknown constants, the set of possible pdfs is usually written as $p_{\mathbf{z}}(\mathbf{z};\mathbf{u})$, where $\mathbf{z}$ denotes the observation and $\mathbf{u}$ denotes the parameter. When the parameters are treated as random, the *joint* pdf for $\mathbf{z}$ and $\mathbf{u}$ is written as $p_{\mathbf{z},\mathbf{u}}(\mathbf{z},\mathbf{u})$, the *a priori* pdf for $\mathbf{u}$ is the marginal density

$$p_{\mathbf{u}}(\mathbf{u}) = \int p_{\mathbf{z},\mathbf{u}}(\mathbf{z},\mathbf{u})d\mathbf{z},$$

and similarly for $\mathbf{z}$, and the *a posteriori density for* $\mathbf{u}$, i.e., the pdf for $\mathbf{u}$ given the values of the observations, is the conditional density

$$p_{\mathbf{u}|\mathbf{z}}(\mathbf{u}|\mathbf{z}) = \frac{p_{\mathbf{z},\mathbf{u}}(\mathbf{z},\mathbf{u})}{p_{\mathbf{z}}(\mathbf{z})} = \frac{p_{\mathbf{z}|\mathbf{u}}(\mathbf{z}|\mathbf{u})p_{\mathbf{u}}(\mathbf{u})}{p_{\mathbf{z}}(\mathbf{z})}.$$

A similar notation is used for expectations of random quantities, e.g., the conditional expectation of a function $f(\mathbf{u})$ of $\mathbf{u}$, given the value taken by $\mathbf{z}$, is written

$$E_{\mathbf{u}|\mathbf{z}}\{\mathbf{f}(\mathbf{u})\} = \int \mathbf{f}(\mathbf{u})p(\mathbf{u}|\mathbf{z})d\mathbf{u}.$$

Sometimes, when it is clear from the context what is intended, the subscript on p will be omitted, as in the previous equation. In this report, the parameter vector will be treated as random. An estimate of $\mathbf{u}$ is a function only of the observations, $\hat{\mathbf{u}}(\mathbf{z})$. The estimation error, $\hat{\mathbf{u}}(\mathbf{z}) - \mathbf{u}$, is a random vector, and so is completely characterized by the joint pdf $p(\mathbf{z}\ \mathbf{u})$. In many cases, it can be argued that the error vector is approximately Gaussian and so is well characterized by its mean value and covariance matrix, or by its *mean squared error matrix*

$$E_{\mathbf{z},\mathbf{u}}\{(\hat{\mathbf{u}}(\mathbf{z}) - \mathbf{u})(\hat{\mathbf{u}}(\mathbf{z}) - \mathbf{u})^T\}.$$

3. THE WW BOUND

Weiss and Weinstein [2,3,4] have recently derived a bound on the mean squared error [1] in estimating random parameters. The bound states that [4, Equation (18)]

$$E_{\mathbf{z},\mathbf{u}}(\hat{\mathbf{u}}-\mathbf{u})(\hat{\mathbf{u}}-\mathbf{u})^{\mathbf{T}} \geq \mathbf{H}\boldsymbol{\Gamma}^{-1}\mathbf{H}^{\mathbf{T}} \triangleq \boldsymbol{\Omega}, \tag{1}$$

where $\hat{\mathbf{u}}^{K\times 1}$ is the vector estimate, $\mathbf{u}^{K\times 1}$ is the true parameter vector, $\mathbf{H}^{K\times T}$ is a matrix, the columns $\mathbf{h}_i$ of which represent perturbations in the true parameter vector $\mathbf{u}$, and $\boldsymbol{\Gamma}^{T\times T}$ is a matrix to be defined below in terms of the likelihood ratio of the observations when the true parameter vector is $\mathbf{u}+\mathbf{h}$ and $\mathbf{u}$, respectively,

$$L(\mathbf{z};\mathbf{u}+\mathbf{h},\mathbf{u}) \triangleq \frac{p(\mathbf{z},\mathbf{u}+\mathbf{h})}{p(\mathbf{z},\mathbf{u})}. \tag{2}$$

The likelihood function is defined only on the set U of values of $\mathbf{u}$ for which $p(\mathbf{z},\mathbf{u})>0$. Define a real random vector $\mathbf{w}$, the ith element of which is given by

$$w_i \triangleq \frac{L^{1/2}(\mathbf{z};\mathbf{u}+\mathbf{h}_j,\mathbf{u})-L^{1/2}(\mathbf{z};\mathbf{u}-\mathbf{h}_j,\mathbf{u})}{E\{L^{1/2}(\mathbf{z};\mathbf{u}+\mathbf{h}_j,\mathbf{u})\}}.$$

The matrix $\boldsymbol{\Gamma}$ is then simply the covariance matrix of $\mathbf{w}$,

$$\boldsymbol{\Gamma} = E\{\mathbf{w}\mathbf{w}^T\}. \tag{3}$$

It is clear from this representation that $\boldsymbol{\Gamma}$ is nonnegative definite; hereafter it is assumed to be positive definite.

The number T and nature of the perturbation vectors in $\mathbf{H}$ are completely arbitrary. Every choice of them produces a lower bound on the mean-squared error matrix. The addition of another perturbation vector can only increase the bound, in the sense that if the bounding matrix is $\boldsymbol{\Omega}$ before and $\boldsymbol{\Omega}'$ after a perturbation vector is added, then $\boldsymbol{\Omega}'-\boldsymbol{\Omega}\geq\mathbf{0}$. A proof of this result is given in Appendix A.

[1] Their bound is actually more general than the one described here. See Weiss and Weinstein [4] for the more general form.

4. THE SINGLE-SOURCE, IDEAL-ARRAY DF PROBLEM

The mathematical model of the array DF problem for narrowband sources that will be used here has been described in many papers (see, e.g., Johnson [5]). A brief description will be given here in order to introduce the notation to be used. Attention is restricted here to the case of a single narrowband source, so that the complex envelopes of the signals received at each of the M array elements at time t form a complex M vector

$$\mathbf{z}(t) = \beta(t)\mathbf{v}(\mathbf{u}) + \mathbf{n}(t), \tag{4}$$

where $\mathbf{u}$ is a real unit vector pointing from the array origin to the source, $\beta(t)$ is the complex amplitude of the incident wavefront at the array origin, $\mathbf{v}(\mathbf{u})$ is a complex M vector, the *array response vector* for direction $\mathbf{u}$, containing the complex signal components received at each of the elements divided by $\beta(t)$, and $\mathbf{n}(t)$ is a complex M vector containing the noise components at each element. The data consists of samples of $\mathbf{z}(t)$ taken at times $(t_1, t_2, \ldots, t_N)$. For brevity, all vector values at time t_n will be indexed simply by n. It is assumed that the time between samples is large enough so that the noise vectors $\mathbf{n}_n$ and also the complex signal amplitudes β_n are statistically independent. The noise vectors $\mathbf{n}_n$ are modeled as circular complex Gaussian random vectors with zero mean and covariance matrix $\mathbf{I}$.[2] The complex signal amplitude β_n is also modeled as a circular complex Gaussian random variable with zero mean and known, constant variance P.[3] With these assumptions, the data vectors $\mathbf{z}_n$ are independent, circular complex Gaussian random vectors with zero mean and common covariance matrix $\mathbf{C} = P\mathbf{v}(\mathbf{u})\mathbf{v}^H(\mathbf{u}) + \mathbf{I}$.

The signal source is assumed to be far away from the array, so that the incident wavefront is essentially a plane wave. If the mth ideal element[4] is located at position $\mathbf{d}_m$ relative to the array origin, its array response $v_m(\mathbf{u})$ is $\exp(jk\mathbf{u}^T\mathbf{d}_m)$, where $k = 2\pi/\lambda$ is the wave number. The array response vector (transposed) is

$$\mathbf{v}^T(\mathbf{u}) = (\exp(jk\mathbf{u}^T\mathbf{d}_1), \exp(jk\mathbf{u}^T\mathbf{d}_2), \ldots, \exp(jk\mathbf{u}^T\mathbf{d}_M)). \tag{5}$$

---

[2] All $2M$ real and imaginary parts of the entries of $\mathbf{n}_n$ are statistically independent and have equal variances of $1/2$.

[3] If the variance P is unknown, it becomes an additional parameter to be estimated, and must be assigned an *a priori* probability density if one wishes to use the WW bound.

[4] An ideal element measures the complex amplitude of the incident wavefront with an ideal omnidirectional pattern.

Note that $||\mathbf{v}(\mathbf{u})||^2 = M$. Because the signal magnitude at each element is $|\beta|$ and the noise variance at each element is unity, P represents the ESNR at each element.

The array is assumed to be planar. This assumption affects the analysis in that it simplifies the form of the derivative with respect to $\mathbf{u}$ of $\mathbf{v}(\mathbf{u})$. This means that the z component of all the vectors $\mathbf{d}_m$ are zero, and so they and the vector $\mathbf{u}$ may be regarded as two-dimensional. It is also assumed that the direction of arrival is confined to the hemisphere above the plane of the array. This eliminates any ambiguity in the z component of the vector $\mathbf{u}$.

There are thus two real parameters to be estimated, the x and y components of $\mathbf{u}$. An *a priori* pdf must therefore be chosen for these variables. To keep the computation of the bound as simple as possible, this function is chosen to be constant on the unit disk, i.e.,

$$p(\mathbf{u}) = p(u_x, u_y) = \begin{cases} 1/\pi & \text{if } u_x^2 + u_y^2 \leq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

This density implies that the marginal probability density of signals in elevation angle is $\sin 2\phi$. If the signal directions were uniformly distributed in solid angle over a hemisphere, the resulting density in elevation angle would be $\cos \phi$, producing more signals near the horizon and less at higher elevation angles. In many applications, this would be a more realistic model; however, it would greatly complicate the computation of the bound.

5. CALCULATION OF THE WW BOUND

The first step in computing the bound is to write down the joint probability density of the measured data and the parameters, which is the product of the *a priori* density on $\mathbf{u}$ and the *conditional* density of the data when $\mathbf{u}$ is known. This conditional density is circular complex Gaussian, with covariance matrix

$$\mathbf{C}(\mathbf{u}) \triangleq E\{\mathbf{z}\mathbf{z}^H\} = \mathbf{I} + P\mathbf{v}(\mathbf{u})\mathbf{v}^H(\mathbf{u}). \tag{7}$$

The inverse of $\mathbf{C}$ is

$$\mathbf{C}^{-1}(\mathbf{u}) = \mathbf{I} - \frac{P}{MP+1}\mathbf{v}(\mathbf{u})\mathbf{v}^H(\mathbf{u}), \tag{8}$$

and its determinant is

$$\det \mathbf{C}(\mathbf{u}) = MP + 1.$$

The joint probability density of the data $\mathbf{z} \triangleq \{\mathbf{z}_n, n = 1, \ldots, N\}$ and the parameters u_x, u_y is

$$p(\mathbf{z}, \mathbf{u}) = p(\mathbf{z}|\mathbf{u})p(\mathbf{u}) = [\pi^M(MP+1)]^{-N} \exp\{-\sum_{n=1}^{N} \mathbf{z}_n^H \mathbf{C}^{-1}(\mathbf{u})\mathbf{z}_n\} p(\mathbf{u}). \tag{9}$$

The likelihood ratios required in the evaluation of the bound are

$$\begin{aligned} L(\mathbf{z}; \mathbf{u}+\mathbf{h}_i, \mathbf{u}) &\triangleq \frac{p(\mathbf{z}, \mathbf{u}+\mathbf{h}_i)}{p(\mathbf{z}, \mathbf{u})} \\ &= \frac{p(\mathbf{u}+\mathbf{h}_i)}{p(\mathbf{u})} \exp\left\{\frac{P}{MP+1} \sum_{n=1}^{N} (|\mathbf{z}_n^H \mathbf{v}(\mathbf{u}+\mathbf{h}_i)|^2 - |\mathbf{z}_n^H \mathbf{v}(\mathbf{u})|^2)\right\}. \end{aligned}$$

Note that these likelihood ratios are defined only on the set U (the unit disk) where $p(\mathbf{u}) > 0$.

All that is needed to compute the $\mathbf{\Gamma}$ matrix [Equation (3)] required for the WW bound is the evaluation of two expectations. The first is

$$
\begin{aligned}
E\{L^{1/2}(\mathbf{z};\mathbf{u}+\mathbf{h}_i,\mathbf{u})\} &= \int_U d\mathbf{u}\, p(\mathbf{u}) \left(\frac{p(\mathbf{u}+\mathbf{h}_i)}{p(\mathbf{u})}\right)^{1/2} \\
&\cdot \int d\mathbf{z} \exp\left\{\frac{1}{2}\frac{P}{MP+1}\sum_{n=1}^{N}(|\mathbf{z}_n^H\mathbf{v}(\mathbf{u}+\mathbf{h}_i)|^2 - |\mathbf{z}_n^H\mathbf{v}(\mathbf{u})|^2)\right\} \\
&\cdot [\pi^M(MP+1)]^{-N} \exp\left\{-\sum_{n=1}^{N}\mathbf{z}_n^H\mathbf{C}^{-1}(\mathbf{u})\mathbf{z}_n\right\}.
\end{aligned}
$$

The last two factors under the $\mathbf{z}$ integral above are of course just $p(\mathbf{z}|\mathbf{u})$. The right side simplifies to

$$
\int_U d\mathbf{u}\, p^{1/2}(\mathbf{u}+\mathbf{h}_i)p^{1/2}(\mathbf{u}) \int d\mathbf{z}\, [\pi^M(MP+1)]^{-N} \exp\{-\sum_{n=1}^{N}\mathbf{z}_n^H\mathbf{Q}\mathbf{z}_n\}, \tag{10}
$$

where

$$
\mathbf{Q} = \mathbf{I} - \frac{1}{2}\frac{P}{MP+1}(\mathbf{v}(\mathbf{u}+\mathbf{h}_i)\mathbf{v}^H(\mathbf{u}+\mathbf{h}_i) + \mathbf{v}(\mathbf{u})\mathbf{v}^H(\mathbf{u})). \tag{11}
$$

The $\mathbf{z}$ integration reduces to a product of N integrals of the form

$$
[\pi^M(MP+1)]^{-1} \int d\mathbf{z}_n \exp\{-\mathbf{z}_n^H\mathbf{Q}\mathbf{z}_n\} = [(MP+1)\det\mathbf{Q}]^{-1}.
$$

It is shown in Appendix C that

$$
\begin{aligned}
(MP+1)\det\mathbf{Q} &= 1 + \frac{(MP)^2}{4(MP+1)}(1 - |\mathbf{v}^H(\mathbf{u})\mathbf{v}(\mathbf{u}+\mathbf{h}_i)|^2/M^2) \\
&= 1 + \frac{(MP)^2}{4(MP+1)}(1 - |\mathbf{v}^H(\mathbf{0})\mathbf{v}(\mathbf{h}_i)|^2/M^2),
\end{aligned} \tag{12}
$$

where $\mathbf{v}(\mathbf{0}) = \mathbf{1}$. Note that $\det\mathbf{Q}$ is not a function of $\mathbf{u}$, and so the two integrals in Equation (10) are separable. The final result is

$$
E\{L^{1/2}(\mathbf{z};\mathbf{u}+\mathbf{h}_i,\mathbf{u})\} = [(MP+1)\det\mathbf{Q}]^{-N} \int_U d\mathbf{u}\, p^{1/2}(\mathbf{u}+\mathbf{h}_i)p^{1/2}(\mathbf{u}). \tag{13}
$$

The second required expectation is

$$E\{L^{1/2}(\mathbf{z};\mathbf{u}+\mathbf{h}_i,\mathbf{u})L^{1/2}(\mathbf{z};\mathbf{u}+\mathbf{h}_j,\mathbf{u})\} =$$

$$\int_U d\mathbf{u}\, p(\mathbf{u}) \left(\frac{p(\mathbf{u}+\mathbf{h}_i)}{p(\mathbf{u})}\right)^{1/2} \left(\frac{p(\mathbf{u}+\mathbf{h}_j)}{p(\mathbf{u})}\right)^{1/2}$$

$$\cdot[\pi^M(MP+1)]^{-N} \int d\mathbf{z}\, \exp\left\{-\sum_{n=1}^{N} \mathbf{z}_n^H \mathbf{C}^{-1}(\mathbf{u})\mathbf{z}_n\right\}$$

$$\cdot \exp\left\{\frac{1}{2}\frac{P}{MP+1}\sum_{n=1}^{N}(|\mathbf{z}_n^H\mathbf{v}(\mathbf{u}+\mathbf{h}_i)|^2 + |\mathbf{z}_n^H\mathbf{v}(\mathbf{u}+\mathbf{h}_j)|^2 - 2|\mathbf{z}_n^H\mathbf{v}(\mathbf{u})|^2)\right\},$$

which reduces to [compare Equation (10)]

$$\int_U d\mathbf{u}\, p^{1/2}(\mathbf{u}+\mathbf{h}_i)p^{1/2}(\mathbf{u}+\mathbf{h}_j) \int d\mathbf{z}\, [\pi^M(MP+1)]^{-N} \exp\{-\sum_{n=1}^{N} \mathbf{z}_n^H \mathbf{R}\mathbf{z}_n\}, \qquad (14)$$

where [compare Equation (11)]

$$\mathbf{R} = \mathbf{I} - \frac{1}{2}\frac{P}{MP+1}(\mathbf{v}(\mathbf{u}+\mathbf{h}_i)\mathbf{v}^H(\mathbf{u}+\mathbf{h}_i) + \mathbf{v}(\mathbf{u}+\mathbf{h}_j)\mathbf{v}^H(\mathbf{u}+\mathbf{h}_j)). \qquad (15)$$

The same sequence of steps that produced Equation (13) now leads to

$$E\{L^{1/2}(\mathbf{z};\mathbf{u}+\mathbf{h}_i,\mathbf{u})L^{1/2}(\mathbf{z};\mathbf{u}+\mathbf{h}_j,\mathbf{u})\} =$$

$$[(MP+1)\det\mathbf{R}]^{-N} \int_U d\mathbf{u}\, p^{1/2}(\mathbf{u}+\mathbf{h}_i)p^{1/2}(\mathbf{u}+\mathbf{h}_j) \qquad (16)$$

where [compare Equation (12)]

$$\begin{aligned}(MP+1)\det\mathbf{R} &= 1 + \frac{(MP)^2}{4(MP+1)}(1 - |\mathbf{v}^H(\mathbf{u}+\mathbf{h}_i)\mathbf{v}(\mathbf{u}+\mathbf{h}_j)|^2/M^2) \\ &= 1 + \frac{(MP)^2}{4(MP+1)}(1 - |\mathbf{v}^H(\mathbf{h}_i)\mathbf{v}(\mathbf{h}_j)|^2/M^2). \qquad (17)\end{aligned}$$

Again $\det\mathbf{R}$ is not a function of $\mathbf{u}$, and so the two integrals in Equation (14) are separable, yielding Equation (16). Note that the determinant $\det\mathbf{Q}$ given by Equation (12) is just a special case of Equation (17).

In order to obtain a compact expression for the Γ matrix that defines the WW bound, it is convenient to define the functions

$$A_N(\mathbf{h}_i, \mathbf{h}_j) = [(MP+1)\det \mathbf{R}]^{-N}, \tag{18}$$

$$B(\mathbf{h}_i, \mathbf{h}_j) = \int_U d\mathbf{u}\, p^{1/2}(\mathbf{u}+\mathbf{h}_i) p^{1/2}(\mathbf{u}+\mathbf{h}_j), \tag{19}$$

and their product

$$C(\mathbf{h}_i, \mathbf{h}_j) = A_N(\mathbf{h}_i, \mathbf{h}_j) B(\mathbf{h}_i, \mathbf{h}_j).$$

Note that only A_N depends on the signal-to-noise ratio P and the number of samples N. Note also that because the domain of integration in Equation (19) is the set U, the region of support in Equation (19) is the intersection of three unit disks, not two (see Appendix D).

In terms of these functions, the ijth element of the Γ matrix of Equation (3) takes the form

$$\Gamma_{ij} = \frac{C(\mathbf{h}_i, \mathbf{h}_j) + C(-\mathbf{h}_i, -\mathbf{h}_j) - C(\mathbf{h}_i, -\mathbf{h}_j) - C(-\mathbf{h}_i, \mathbf{h}_j)}{C(\mathbf{h}_i, \mathbf{0}) C(\mathbf{h}_j, \mathbf{0})}. \tag{20}$$

It is clear from their definitions that A_N and B have the common property

$$A_N(\mathbf{h}_i, \mathbf{h}_j) = A_N(\mathbf{h}_j, \mathbf{h}_i),$$

$$B(\mathbf{h}_i, \mathbf{h}_j) = B(\mathbf{h}_j, \mathbf{h}_i),$$

and that A_N is an even function in the sense that

$$A_N(-\mathbf{h}_i, -\mathbf{h}_j) = A_N(\mathbf{h}_i, \mathbf{h}_j).$$

If, as has been assumed here, the parameter set U has the property that if $\mathbf{u}$ is in U, $-\mathbf{u}$ is also in U, and that the *a priori* density $p(\mathbf{u})$ is an even function, i.e., $p(-\mathbf{u}) = p(\mathbf{u})$, then Equation (20) simplifies to

$$\Gamma_{ij} = \frac{2[C(\mathbf{h}_i, \mathbf{h}_j) - C(\mathbf{h}_i, -\mathbf{h}_j)]}{C(\mathbf{h}_i, \mathbf{0}) C(\mathbf{h}_j, \mathbf{0})}.$$

6. A SCALAR MEASURE OF DF ACCURACY

A planar array can estimate two real angular parameters (here u_x, u_y) that define a unique direction in a three-dimensional hemisphere. The DF error is thus a two-vector having a two-dimensional probability density. Different array geometries will produce distributions having different shapes. In order to compare two such arrays, their distributions must be characterized by a single real number.

A convenient choice is the area of the *ellipse of concentration* of the error distribution (e.g., see Cramer [6, p.283]). A probability density for the error uniformly distributed over the ellipse of concentration would have the same mean squared error matrix as the actual *a posteriori* density. This area of the ellipse of concentration is given by $4\pi\sqrt{\det \boldsymbol{\Xi}}$, where $\boldsymbol{\Xi}$ is the mean squared error matrix of the distribution. The array having the smallest area of concentration is deemed the best. It is convenient to plot this error area on a logarithmic scale.

In the limit as ESNR goes to zero, the observations provide no information and the mean squared error matrix becomes that of the *a priori* density. This has $\boldsymbol{\Xi} = \mathbf{I}/4$, so that $\det \boldsymbol{\Xi} = 1/16$, and its ellipse of concentration has area π. This is the value the scalar measure should approach for low ESNR.

7. THE CR FACTOR OF AN ARRAY

General expressions for the Fisher information matrix **F** for *nonrandom* parameter estimation are given in DeLong [7] for the case of a multiple signal scenario in which each signal has a single unknown directional parameter. Both deterministic and random Gaussian signal models are considered. In the deterministic model, the complex amplitudes may be either known or unknown. In the Gaussian model, the covariance matrix of the signal complex amplitudes may be either known, or unknown but diagonal (uncorrelated signals). The Gaussian, known covariance bounds derived from this information matrix (assuming unbiased estimates) exhibit unusual behavior in the multiple signal case, which has been verified by simulation of estimation algorithms in White [8].

The Fisher matrix of DeLong [7, Equation (2.242)], for the Gaussian-signal, known-covariance problem, is extended to signals with an arbitrary number of directional parameters (e.g., the planar array problem) in Appendix F. When specialized to the case of a planar array of omnidirectional elements, it becomes

$$\mathbf{F} = 2k^2 N \frac{MP^2}{1+MP}\left[\mathbf{D}^T(\mathbf{I} - \frac{1}{M}\mathbf{1}\mathbf{1}^T)\mathbf{D}\right], \tag{21}$$

where

$$\mathbf{D}^T = \begin{pmatrix} x_1 & x_2 & \cdots & x_M \\ y_1 & y_2 & \cdots & y_M \end{pmatrix} \tag{22}$$

is a matrix containing the x, y positions of the elements of the array in some arbitrary coordinate system. Note that **F** is not a function of **u**. This fact will be quite useful below. For Gaussian signals, the quantity $\mathbf{F}^{-1}$ is a lower bound on the covariance matrix of any unbiased estimate of **u**. It is referred to hereafter as the *unknown parameter* CR bound.

For large signal-to-noise ratios, the regime of interest below, Equation (21) becomes

$$\mathbf{F} = 2k^2 NP\left[\mathbf{D}^T(\mathbf{I} - \frac{1}{M}\mathbf{1}\mathbf{1}^T)\mathbf{D}\right] \triangleq 2k^2 NP\mathbf{A}.$$

The array configuration appears in this expression entirely by way of the matrix **A**. This matrix is simpler than it appears at first glance. If the origin of the array coordinate system is moved to the center of mass of the element positions (so that $\mathbf{D}^T\mathbf{1} = \mathbf{0}$), it becomes simply

$$\mathbf{A} = \mathbf{D}^T\mathbf{D} = \sum_{m=1}^{M} \begin{bmatrix} x_m^2 & x_m y_m \\ x_m y_m & y_m^2 \end{bmatrix}.$$

In the limit of high ESNR, it has been found empirically that the WW lower bound on the mean squared error matrix, for the problem of interest here, becomes essentially equal to the unknown parameter CR bound when the test points are chosen properly (see Section 8). That is, at high ESNR, both bounds say that

$$\boldsymbol{\Xi} \geq \mathbf{F}^{-1}. \tag{23}$$

The area of the ellipse of concentration [6] associated with the error matrix $\boldsymbol{\Xi}$ is $4\pi\sqrt{\det \boldsymbol{\Xi}}$. Because $\det \boldsymbol{\Xi} \geq (\det \mathbf{F})^{-1}$, a lower bound on this area is therefore

$$4\pi\sqrt{\det \boldsymbol{\Xi}} \geq (2\pi N P)^{-1}\lambda^2(\det \mathbf{A})^{-1/2}. \tag{24}$$

Thus, for high ESNR, both the WW and CR approaches yield the DF error bound given by Equation (24). The *CR factor* of the array is defined to be the quantity $(\det \mathbf{A})^{1/4}$, which has the dimension of length. *Arrays having equal CR factors and operating at the same frequency have equal DF accuracy bounds at high signal-to-noise ratios.* It is sometimes convenient to include the factor λ^{-1} in the definition of CR factor. When this is done, the CR factor is said to have units of wavelength. The caveat about operating frequency can be dropped from the above statement if the CR factors of the arrays are equal in wavelengths.

A similar quantity for line arrays, $L = [\frac{1}{M}\sum_{m=1}^{M} x_m^2]^{1/2}$, was developed in Evans et al. [9], where it was called *root mean square (RMS) array length.* However, this parameter has the undesirable property that removing an element can cause it to increase. It is easy to show that the CR factor defined here must decrease when elements are removed.

Although the CR factor has the dimension of length, an array with constant CR factor must shrink as elements are added. For this reason, the terms *length* and *aperture* are not used to describe it.

8. CHOICE OF TEST POINTS

Intuition suggests and simulation confirms that the abrupt increase of the estimation error as signal-to-noise ratio decreases occurs because noise causes a spurious sidelobe peak to exceed the true peak of the likelihood function (or whatever function is being maximized by the estimator). The unknown parameter CR bound takes into account variations of the pdf only in the immediate neighborhood of the true parameter value [see Equation (F.1)]. Thus the performance threshold is not well-predicted by this bound. As will be seen below, the test points of the WW bound can be chosen so as to predict the performance threshold quite well.

Another well known bound, the Barankin bound, is very similar in form to the WW bound and also involves the use of test points. McAulay and Seidman [10] applied the Barankin bound to a pulse-position modulation communications problem and suggested choosing the test points to be the locations of the maximum sidelobes of the objective function (the matched filter output).

It is clear that the height of a sidelobe is not the only factor affecting the bound. If the sidelobe is very distant in arrival angle from the mainlobe, the probability that the true direction of arrival is such that the sidelobe is within "visible space," i.e., that it represents another valid direction of arrival, is very small. Thus such a sidelobe contributes relatively little to the overall DF error. Sidelobes near the mainlobe have a greater probability of causing an error, but the size of the error is smaller.

Weiss and Weinstein [2,3,4] assert that if only incremental test points

$$\mathbf{h}_x = \begin{pmatrix} \epsilon \\ 0 \end{pmatrix}, \quad \mathbf{h}_y = \begin{pmatrix} 0 \\ \epsilon \end{pmatrix}$$

are used, the WW bound becomes the *random parameter* CR bound derived by Van Trees [1, pp.72,84] as $\epsilon \rightarrow 0$. It is shown in Appendix B, for a simple one-parameter problem, that this is not the case when the joint pdf $p(\mathbf{z}, u)$ has compact support in u and is not zero at the end points. The WW bound instead approaches 0 for this case. This behavior is also exhibited by the computer program used to compute the bound for the planar array problem of interest here.

Despite this fact, it is found empirically that there exists, for every instance so far examined, an optimal value of ϵ which yields the tightest bound. Furthermore, the bound so obtained is extremely close to that obtained from the *unknown parameter* CR bound. The empirically-determined equation for ϵ is

$$10\log_{10}\epsilon = \min[0, 0.5862 * (40 - 10\log_{10}(NML^2)) - \mathrm{ESNR(dB)} - 30],$$

where N, M and L are the number of looks, number of elements, and CR factor in wavelengths, respectively, of the array. Use of this formula gives a two-point WW bound, which follows the unknown parameter CR bound closely for ESNRs ranging from -10 to 40 dB.

Experience shows that when the test points of the WW bound are placed at the locations of the highest sidelobes of the array pattern, only a few test points are needed. The addition of more test points has very little effect on the bound. An example of this is shown in Figure 1 for a

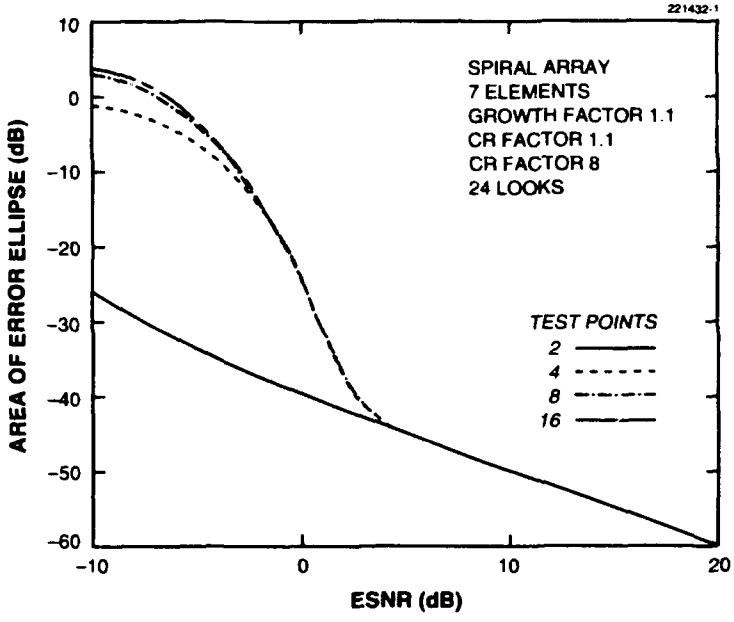

Figure 1. Effect of number of test points on the WW bound.

7-element spiral array having a growth factor of 1.1 and a CR factor of 8 wavelengths, processing 24 looks of array data. (The spiral array geometry and parameters are described more fully in Section 11.) The incremental test points are the first two used; they produce a smooth curve without a break point, essentially the unknown parameter CR bound. Addition of two more test points at the two highest sidelobe locations produces the threshold behavior, and adding more test points merely raises the bound slightly in the uninteresting ESNR region below threshold. At high ESNR, the WW bounds agree with the unknown parameter CR bound, as discussed in Section 7, regardless of the number of test points (≥ 2) used.

The next section discusses the array pattern and the selection of sidelobes in more detail.

9. COMPUTING THE ARRAY POWER PATTERN

The *array power pattern function* of an array is defined to be the squared magnitude of its response, when steered in direction $\mathbf{u}_1$, to a signal incident from direction $\mathbf{u}_2$, i.e.,

$$g(\mathbf{u}_1, \mathbf{u}_2) = |\frac{1}{M}\mathbf{x}(\mathbf{u}_1)^H \mathbf{x}(\mathbf{u}_2)|^2, \tag{25}$$

where $\mathbf{x}(\mathbf{u})$ is the array response vector for direction $\mathbf{u}$. When the array is composed of ideal, omnidirectional elements, the response vectors $\mathbf{x}(\mathbf{u})$ become the vectors $\mathbf{v}(\mathbf{u})$ given by Equation (5). It is then true that

$$g(\mathbf{u}_1, \mathbf{u}_2) = g(\mathbf{0}, \mathbf{u}_2 - \mathbf{u}_1) \triangleq g(\mathbf{h}), \tag{26}$$

where $\mathbf{h} = \mathbf{u}_2 - \mathbf{u}_1$. An array is characterized by computing $g(\mathbf{u}_1, \mathbf{u}_2)$ over all possible pairs of unit vectors in a hemisphere. From Equation (26), it is seen that this is equivalent to steering the array to zenith ($\mathbf{u} = \mathbf{0}$) and computing over all possible *differences* of pairs of unit vectors in a hemisphere. It is clear that the maximum length of the difference of two unit vectors is 2.

The global maximum of the array pattern occurs at $\mathbf{h} = \mathbf{0}$. Secondary maxima of the pattern (which may sometimes be equally large) are called the sidelobes of the array pattern. The location of a sidelobe is specified by a vector $\mathbf{h}$. For a general array, $\mathbf{h}$ is three dimensional, but for planar arrays (the case considered here) its component normal to the plane is of no consequence. The array pattern is then a function of two real variables, the x and y components of $\mathbf{h}$.

A list of sidelobe values arranged in descending order, accompanied by their locations ($\mathbf{h}$ values), defines the array as far as the WW bound is concerned. When the array is steered to zenith, the $\mathbf{h}$ values define the sidelobe locations directly; otherwise, they define the difference vector between the steering direction and the sidelobe direction.

While the array pattern is translation-invariant, the effect of a sidelobe is not translation-invariant. When the array is steered to zenith, the only sidelobes that are important are those lying within the unit circle in the $\mathbf{h}$ plane. This is sometimes called *visible space*, because it contains all the physically possible directions of arrival. When the array is steered to direction $\mathbf{u}_0$, the main peak moves there, and the sidelobes move with it. However, visible space is still the unit circle. Thus, as the beam is steered in different directions, some sidelobes vanish from visible space, while others enter. DF accuracy is thus very much a function of steering direction.

The WW bound overcomes this complexity by averaging over steering direction, using an assumed *a priori* probability density for directions of arrival, as described in Section 4. A fixed set of test points is used; whether or not they make a contribution to the bound is determined by the factors $B(\mathbf{h}_i, \mathbf{h}_j)$ defined by Equation (19).

When computing the array pattern, one must choose a wavelength. This should be the smallest wavelength of interest. As the wavelength is increased, the sidelobes of the array pattern move radially outward in direct proportion. The sidelobe list obtained for a chosen wavelength can thus be used as a source of test points for the WW bound computation at any longer wavelength.

10. VALIDATION OF THE WW BOUND

A single-signal DF simulation was carried out in order to evaluate the tightness of the bound. The DF algorithm used in the simulation was the maximum likelihood estimator, discussed in Appendix E. The array configuration was a 5-element spiral[5] with a growth factor of $2^{1/5}$ and a CR factor of 4 wavelengths.

Each simulation point is the result of averaging 1000 Monte Carlo trials. On each trial, a point (u_x, u_y) was selected at random from the unit disk, representing the true direction of arrival. Using these values in the data model of Equation (4), $N = 24$ data vectors (looks) were generated and used to compute the quantity

$$f(\hat{\mathbf{u}}) \triangleq \sum_{n=1}^{N} |\mathbf{z}_n^H \mathbf{v}(\hat{\mathbf{u}})|^2,$$

which is monotonically related to the likelihood function (see Appendix E). The maximum value of this function on the unit disk $|\hat{\mathbf{u}}| \leq 1$ was found by a two-dimensional search.

The results are shown in Figure 2. Note that two results are given for ESNRs of 15 and 17 dB. This is to emphasize the fact that, near threshold, the increase in the error is due to a few occurrences of a rare event (confusion of the main response with a sidelobe response). In the first set of 1000 trials, this event did not occur. A closer match would presumably have been obtained by conducting more trials, but the intensive computation required by the two-dimensional search makes this impractical. The implication is that simulation is not an efficient way to determine threshold ESNR. In view of the difficulties inherent in such a simulation, the bound may be said to accurately predict the threshold ESNR of this array.

[5]For a more complete description of the spiral configuration, see Section 11.

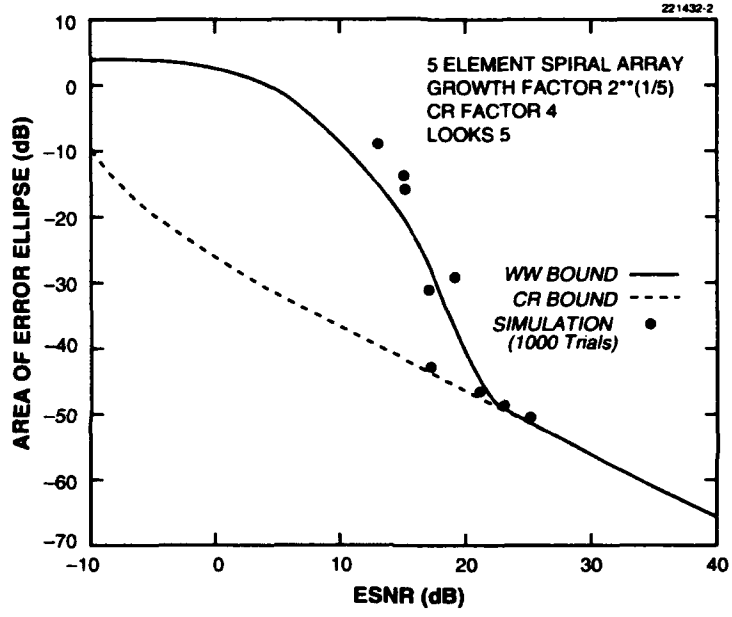

Figure 2 Comparison of simulation results with WW bound.

11. COMPARISON OF THREE ARRAY GEOMETRIES

This section shows how the WW bound can be used to compare the DF accuracy of different array configurations. The three arrays to be studied have the same CR factor and therefore the same accuracy bounds at sufficiently high signal-to-noise ratios. The best array is then the one with the lowest threshold signal-to-noise ratio.

The arrays to be compared all have seven elements. The first is a circular array with elements evenly spaced in angle. For a CR factor of 1 wavelength, this array has a radius of 0.535 wavelength.

The second array is a spiral array. Its elements are on the same radials as the circular array, but their distance from the center grows geometrically by the factor b. This parameter will be called the *growth factor* of the array. In this example, the growth factor is 1.1. For a CR factor of 1 wavelength, the first element is at a distance of 0.391 wavelength from the center.

The third array has the shape of the letter L. There is an element at the intersection and three along each arm. Along each of the two arms, the element positions are geometric, i.e., r, br, b^2r. For radial arrays such as this, the parameter b will be called the *growth factor*. In this example, b is 1.7. For a CR factor of 1 wavelength, the distance r is 0.840 wavelength. Although several growth factors were tried for the spiral and L arrays, no further attempt was made to optimize these parameters.

The three array geometries are shown in Figure 3.

The CR factor of an array is increased by expanding the array or, equivalently, increasing the frequency of operation. Bounds on DF error for the circular, spiral, and L-shaped arrays, respectively, are shown as a function of element signal-to-noise ratio in Figures 4 to 6 for various CR factors. The threshold ESNR[6] for each array is plotted as a function of CR factor in Figure 7.

The circular array clearly has the lowest threshold ESNR, while the spiral is somewhat better than the L for most CR factors. This Figure also illustrates the general principle that the threshold ESNR is a nondecreasing function of CR factor. The reason for this is that increasing the size of an array, and hence its CR factor, causes additional sidelobes to be drawn into visible space.

It must be emphasized that the purpose of this section is not to advocate the circular, or any other, array geometry as being generally superior. It is simply to compare three particular arrays on an equal basis using the WW bound as a tool.

---

[6]Threshold ESNR is defined computationally as that value of ESNR at which the multitest-point WW bound has diverged from the two-test-point version by 0.25 dB. It is computed to a precision of 0.2 dB in ESNR.

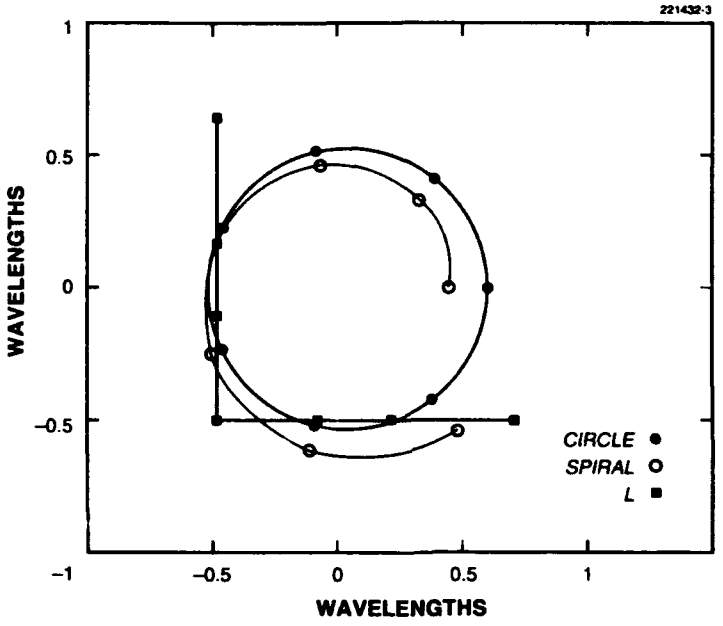

Figure 3. Circular, spiral, and L-shaped arrays having CR factors of unity.

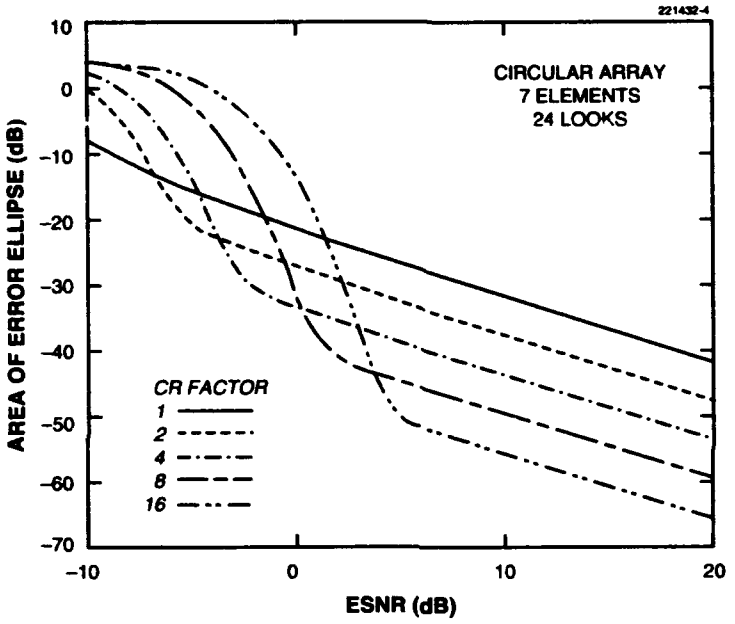

Figure 4. Bounds on DF error for circular array for various CR factors.

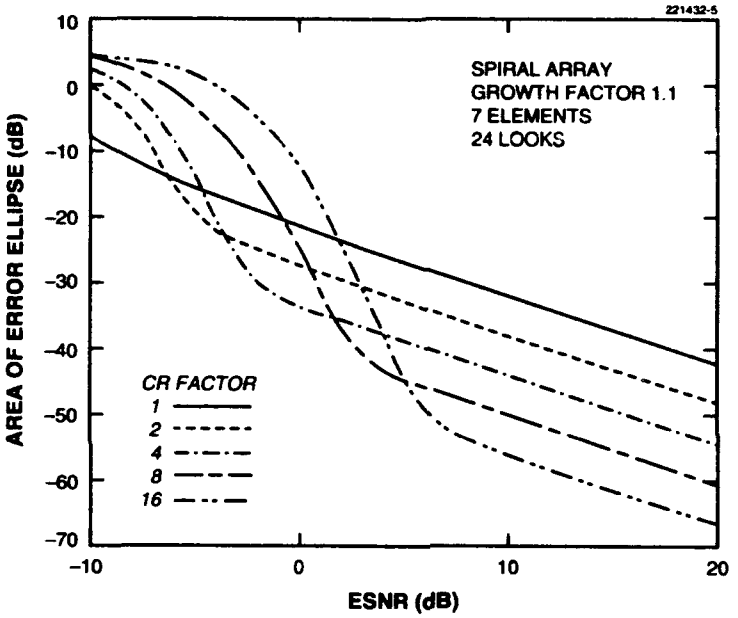

Figure 5. Bounds on DF error for spiral array for various CR factors.

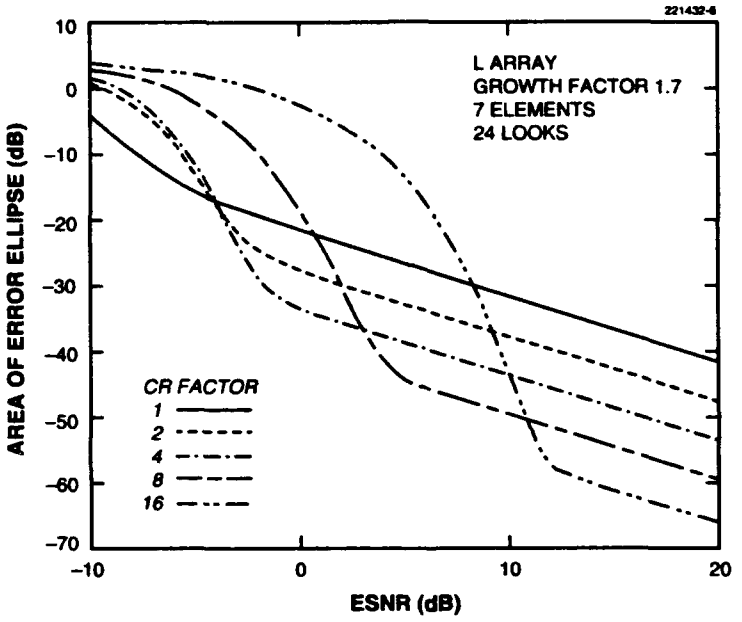

Figure 6. Bounds on DF error for L-shaped array for various CR factors.

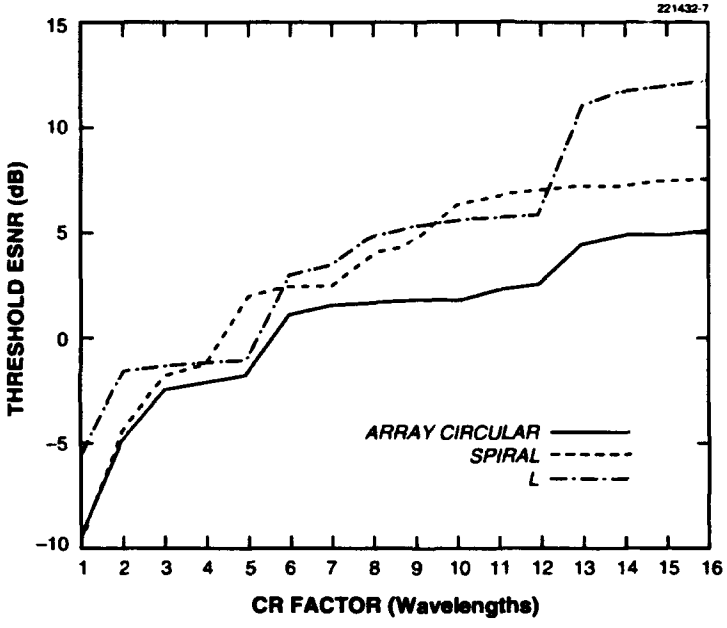

Figure 7. Threshold ESNR behavior of the three arrays vs. CR factor.

12. EXTENSION TO SWITCHED SUBARRAYS

One possible solution to the problem of doing reliable DF with a small number of receiver channels is to switch the available channels among a larger number of antenna elements. The database from which the DF is derived then consists of several looks taken from each of several subarrays. The various subarray patterns have high sidelobes in different directions, and these tend to "average out," producing a pattern not unlike that of the full array.

Calculation of the WW bound for switched subarrays requires only minimal changes in the previous analysis. In Equation (9), the sum that appears in the exponential is replaced by a sum of sums, one for each subarray, and the first factor becomes a product of different factors. The covariance matrix **C** is different for each subarray.

The effect of this change on the previous analysis is simply to replace Equation (18) with a product of terms, each having the same form as Equation (18).

$$A_N(\mathbf{h}_i, \mathbf{h}_j) = \prod_{s=1}^{S} [(M_s P + 1) \det \mathbf{R}_s]^{-N_s},$$

where M_s is the number of elements, $\mathbf{R}_s$ is the matrix of Equation (15), and N_s is the number of looks, all for the sth subarray.

The test points in the switched subarray case are chosen to be the highest sidelobes of the *switched subarray pattern*, which is a weighted sum of the subarray patterns. The mathematical definition of the switched subarray pattern and the rationale behind this definition are given in Appendix E.

Figure 8 presents the results of a Monte Carlo simulation that validates the WW bound calculation for switched subarrays. The simulated array is a 7-element spiral with a growth factor of 1.1 and a CR factor of 4. Three 5-element subarrays are used, each providing eight looks. The three subarrays are comprised of elements (1,2,4,6,7), (1,3,4,5,7), and (2,3,4,5,6), element 1 being innermost. As before, each simulation point represents an average of 1000 Monte Carlo trials. At each trial, the true direction cosines u_x, u_y are chosen at random from the unit disk. The calculated bound again is seen to give a reasonably accurate prediction of the threshold ESNR.

Figure 9 shows the predicted DF accuracy of the switched subarrays for various CR factors (the CR factor is that of the full array). This can be compared directly with the unswitched array accuracy shown in Figure 5. In all cases, there is a loss of about 2 dB in accuracy at high ESNR (because only 16 looks are available from most elements instead of 24), and the threshold ESNR is increased by approximately 2 dB. Similar results are obtained for the circular and L-shaped arrays.

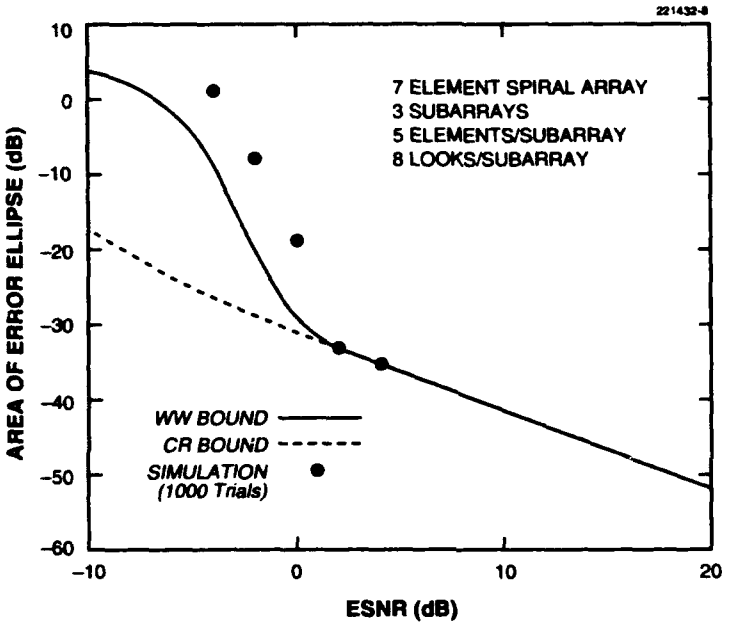

Figure 8. Comparison of simulation results for switched spiral array with WW bound.

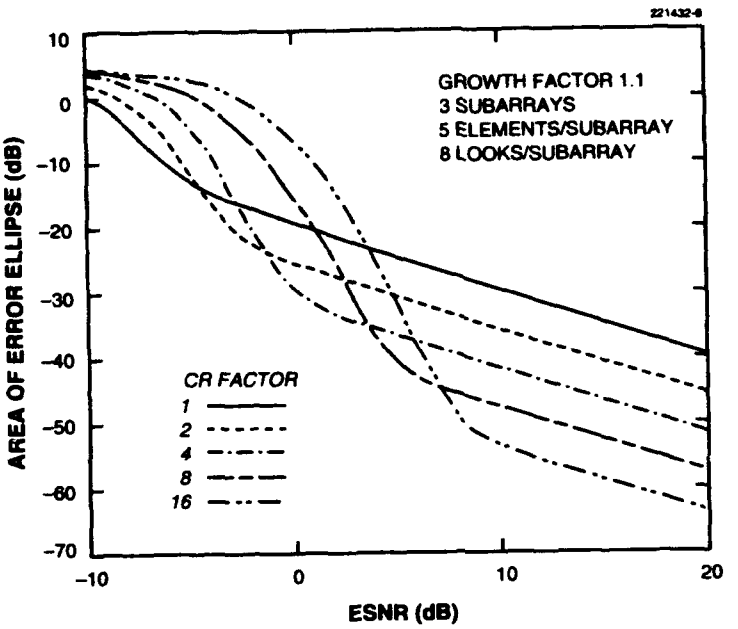

Figure 9. DF error bounds for switched spiral array.

13. SUMMARY

A method has been developed for comparing sparse array configurations of ideal elements based on DF performance. Arrays to be compared are sized so that each has the same DF error bound at high signal-to-noise ratio. They are then ranked in terms of their threshold signal-to-noise ratios, the lowest being deemed the best. The method makes use of the WW bound as it applies to the array DF problem. This bound accurately predicts the threshold signal-to-noise ratio of an ideal array, without the need for computer simulation.

As an example of its use, three array configurations, a circle, a spiral, and an L, are compared.

The WW bound can also be computed for a slightly more general situation in which data is available sequentially in time from several subarrays. In this way it is possible to collect data from a large array with a smaller number of multiplexed receiver channels. As an example, it is shown that the spiral array already examined can be operated in this mode with only a small DF performance degradation. Similar results are obtained with the circular and L arrays.

It is almost certainly true that the threshold ESNR of a real array is determined by calibration errors, rather than by the noise and small sample effects that the bound takes into account. One should not expect a real array to achieve the degree of DF accuracy predicted by the bound. However, it is hoped that if one array is better than another in the sense of this report, and they are implemented with equal attention to calibration issues, the ordering will be preserved.

APPENDIX A
PROOF THAT EACH ADDITIONAL PERTURBATION VECTOR INCREASES THE BOUND

The following easily verified factorization of a partitioned matrix is valid whenever $\mathbf{A}$ is invertible.

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{CA}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}.$$

Assuming that $\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$ is also invertible, the inverse of this factorization is

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{I} & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{CA}^{-1} & \mathbf{I} \end{pmatrix}.$$

When a perturbation vector is added, the $\mathbf{H}$ matrix becomes $(\mathbf{H}\ \mathbf{h})$, while $\mathbf{\Gamma}$ becomes

$$\begin{pmatrix} \mathbf{\Gamma} & \mathbf{g} \\ \mathbf{g}^T & a \end{pmatrix}.$$

Use of the above factorization for this matrix gives for the augmented bound matrix

$$\begin{aligned} \mathbf{\Omega}' &= (\mathbf{H}\ \mathbf{h}) \begin{pmatrix} \mathbf{\Gamma} & \mathbf{g} \\ \mathbf{g}^T & a \end{pmatrix}^{-1} (\mathbf{H}\ \mathbf{h})^T \\ &= (\mathbf{H}\ \mathbf{v}) \begin{pmatrix} \mathbf{\Gamma}^{-1} & \mathbf{0} \\ \mathbf{0} & (a - \mathbf{g}^T\mathbf{\Gamma}^{-1}\mathbf{g})^{-1} \end{pmatrix} (\mathbf{H}\ \mathbf{v})^T \\ &= \mathbf{H}\mathbf{\Gamma}^{-1}\mathbf{H}^T + c\mathbf{v}\mathbf{v}^T \end{aligned}$$

where

$$\mathbf{v} = \mathbf{h} - \mathbf{H}\mathbf{\Gamma}^{-1}\mathbf{g},$$

and

$$c = (a - \mathbf{g}^T \boldsymbol{\Gamma}^{-1} \mathbf{g})^{-1}$$

is a nonnegative number because the expanded $\boldsymbol{\Gamma}$ matrix must be nonnegative definite. It is now evident that

$$\mathbf{x}^T \boldsymbol{\Omega}' \mathbf{x} \geq \mathbf{x}^T \boldsymbol{\Omega} \mathbf{x}$$

for any vector $\mathbf{x}$, and so $\boldsymbol{\Omega}' \geq \boldsymbol{\Omega}$.

APPENDIX B
ON THE CONVERGENCE OF THE WW BOUND TO THE RANDOM PARAMETER CR BOUND

As mentioned in Section 8, it has been claimed that if only incremental test points are used, the WW bound becomes the *random parameter* CR bound in the limit as $\epsilon \to 0$. The example given here shows that this is not always the case.

For simplicity, attention will be restricted here to the one-parameter case, so that the bound is a scalar quantity. In this case, for a single incremental test point,

$$\begin{aligned} H\Gamma^{-1}H &= \epsilon^2 \frac{[E\{L^{1/2}(\mathbf{z};u+\epsilon,u)\}]^2}{E\{[L^{1/2}(\mathbf{z};u+\epsilon,u)-L^{1/2}(\mathbf{z};u-\epsilon,u)]^2\}} \\ &= \frac{[\int d\mathbf{z}\int du\ p^{1/2}(\mathbf{z},u+\epsilon)p^{1/2}(\mathbf{z},u)]^2}{\frac{1}{\epsilon^2}\int d\mathbf{z}\int du\ [p^{1/2}(\mathbf{z},u+\epsilon)-p^{1/2}(\mathbf{z},u-\epsilon)]^2}. \end{aligned} \tag{B.1}$$

It must be remembered that L, defined by Equation (2), is defined only for values of u where $p(\mathbf{z},u) > 0$ and so the u integration is limited to this region.

The integral in the numerator approaches unity as ϵ goes to zero (see, e.g., Rudin [11, Thm 9.5]). In the denominator, if the interchange of the u integration and the limit operation can be justified, then clearly the bound approaches

$$\left[4\int d\mathbf{z}\int du\ (\frac{\partial}{\partial u}p^{1/2}(\mathbf{z},u))^2\right]^{-1} = \left[\int d\mathbf{z}\int du\ (\frac{\partial}{\partial u}\ln p(\mathbf{z},u))^2 p(\mathbf{z},u)\right]^{-1}, \tag{B.2}$$

which is the CR bound for random parameter estimation as given by Van Trees [1].

On the other hand, suppose that for all $\mathbf{z}$, $p(\mathbf{z},u)$ is nonzero only on a finite interval $[a \le u \le b]$ and is nonzero and continuous from the left or right as appropriate at the end points. Then the interchange referred to above cannot be justified, as the following argument shows. The u-integrals are then over the finite interval $[a,b]$, and, for $\epsilon > 0$, the denominator of Equation (B.1) is

$$\begin{aligned} &\frac{1}{\epsilon^2}\int d\mathbf{z}\left[\int_a^{a+\epsilon} du\ p(\mathbf{z},u+\epsilon)+\int_{b-\epsilon}^{b} du\ p(\mathbf{z},u-\epsilon)\right] \\ &\quad+\frac{1}{\epsilon^2}\int d\mathbf{z}\int_{a+\epsilon}^{b-\epsilon} du\ [p^{1/2}(\mathbf{z},u+\epsilon)-p^{1/2}(\mathbf{z},u-\epsilon)]^2. \end{aligned}$$

Because $p(\mathbf{z},u)$ is not zero at its end points and is continuous from the right or left, as appropriate, the expression within the large brackets is of order ϵ, and so the limit of the first term above as $\epsilon \downarrow 0$ is $+\infty$. The remaining term is positive, so the denominator is $+\infty$ and the bound goes to zero. On

the other hand, Equation (B.2) may well give a finite result, as for example if $\mathbf{z}$ is a complex scalar and $p(z,u)$ is defined on $C^1 \times [0,1]$ by

$$p(z,u) = p(z|u)p(u) = \frac{1}{\pi}\exp\{-(z-u)^2\}(u+\frac{1}{2}).$$

This problem can be avoided by choosing the *a priori* densities so that $p(\mathbf{z},u)$ and its u-derivative are zero at the ends of the interval; however, this may greatly complicate the numerical calculation of the bound.

APPENDIX C
EVALUATION OF A DETERMINANT

The determinant of

$$\mathbf{Q} = \mathbf{I} - c(\mathbf{x}\mathbf{x}^H + \mathbf{y}\mathbf{y}^H),$$

where c is real, can be calculated by doing an eigendecomposition. Because $\mathbf{Q}$ is Hermitian, it has exactly M eigenvectors. $M-2$ of these are orthogonal to the subspace spanned by the vectors $\mathbf{x}, \mathbf{y}$ and have eigenvalue 1. The remaining two must be linear combinations of $\mathbf{x}, \mathbf{y}$.[7] To determine these, solve

$$(\mathbf{I} - c\mathbf{x}\mathbf{x}^H - c\mathbf{y}\mathbf{y}^H)(\mathbf{x} + a\mathbf{y}) = \lambda(\mathbf{x} + a\mathbf{y}),$$

which reduces to

$$(1 - \lambda - c(\mathbf{x}^H\mathbf{x} + a\mathbf{x}^H\mathbf{y}))\mathbf{x} + ((1-\lambda)a - c(\mathbf{y}^H\mathbf{x} + a\mathbf{y}^H\mathbf{y})\mathbf{y}) = 0.$$

The coefficients of $\mathbf{x}$ and $\mathbf{y}$ must both be zero. The value of a that makes the coefficient of $\mathbf{y}$ zero is

$$a = \frac{c\mathbf{y}^H\mathbf{x}}{1 - \lambda - c\mathbf{y}^H\mathbf{y}}.$$

Substitution of this result into the remaining coefficient yields a quadratic equation for $1-\lambda$,

$$(1-\lambda)^2 - c(\mathbf{x}^H\mathbf{x} + \mathbf{y}^H\mathbf{y})(1-\lambda) + c^2(\mathbf{x}^H\mathbf{x}\mathbf{y}^H\mathbf{y} - |\mathbf{x}^H\mathbf{y}|)^2) = 0,$$

with the solution

$$\lambda = 1 - \frac{c}{2}\left(\mathbf{x}^H\mathbf{x} + \mathbf{y}^H\mathbf{y} \pm \sqrt{(\mathbf{x}^H\mathbf{x} - \mathbf{y}^H\mathbf{y})^2 + 4|\mathbf{x}^H\mathbf{y}|^2}\ \right).$$

The value of the determinant of $\mathbf{Q}$ is the product of the eigenvalues.

---

[7] If $\mathbf{x}$ and $\mathbf{y}$ are colinear, the problem reduces to a simpler one.

In the problem addressed by this report, the vectors $\mathbf{x}, \mathbf{y}$ have equal length $\sqrt{M}$, so that the previous equation reduces to

$$\lambda = 1 - c(M \pm |\mathbf{x}^H \mathbf{y}|),$$

and the determinant is

$$\det \mathbf{Q} = (1 - Mc)^2 - (c|\mathbf{x}^H \mathbf{y}|)^2.$$

APPENDIX D
CALCULATION OF $B(\mathbf{h}_i, \mathbf{h}_j)$

The calculation of $B(\mathbf{h}_i, \mathbf{h}_j)$ involves evaluating the integral in Equation (19) with $p(\mathbf{u})$ given by Equation (6). The region of integration U is a disk of unit radius centered at the origin, and the factors in the integrand are nonzero only on disks of unit radius centered at $-\mathbf{h}_i, -\mathbf{h}_j$, respectively. The fundamental problem, then, is to find the area of intersection of three unit disks.

If two disks of unit radius intersect, the area of intersection is the union of two circular segments (see Figure D-2). The inclusion of a third disk of unit radius can have four possible results:

(a) Same area of intersection as above [Figure D-2(a)].

(b) Reduced area of intersection still composed of two circular segments [Figure D-2(b)].

(c) Reduced area of intersection bounded by *three* circular arcs, which is the union of three circular segments and a triangle [Figure D-2(c)].

(d) No intersection.

The area of a circular segment (see Figure D-3) that subtends angle θ is

$$A_{seg} = \frac{\theta - \sin\theta}{2}. \tag{D.1}$$

The angle θ can be expressed in terms of the coordinates of the corners of the circular segment as

$$\theta = 2\arcsin(d/2), \qquad d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}. \tag{D.2}$$

The area of a triangle with vertices at $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ is

$$A_{tri} = \frac{1}{2}\left|\det \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix}\right|. \tag{D.3}$$

These formulas make it possible to evaluate Equation (19) without resorting to numerical integration.

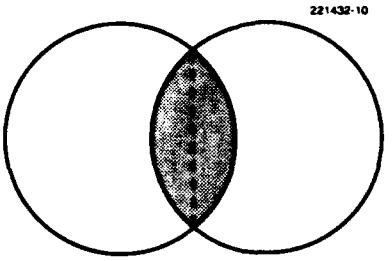

Figure D-1. Intersection of two unit disks.

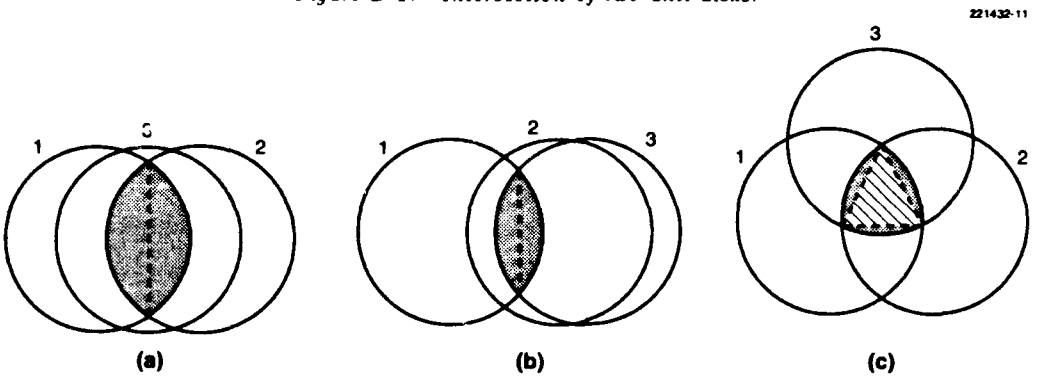

Figure D-2. Possible intersections of three unit disks.

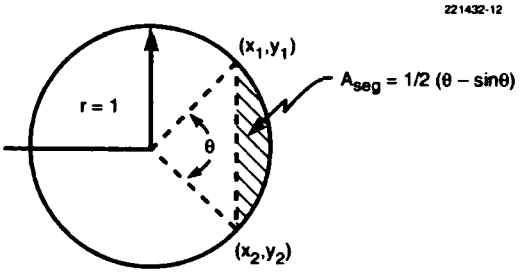

Figure D-3. Area of circular segment.

APPENDIX E
MAXIMUM LIKELIHOOD ESTIMATOR FOR SINGLE-SIGNAL DIRECTION-FINDING WITH NORMAL AND SWITCHED ARRAYS

Let $\mathbf{z}_{s,1}, \mathbf{z}_{s,2}, \ldots, \mathbf{z}_{s,N_s}$ be independent M_s-dimensional vector measurements taken from the sth subarray, $s = 1, \ldots, S$. It is assumed that these measurements are of the form

$$\mathbf{z}_{s,n} = \beta_{s,n}\mathbf{v}_s(\mathbf{u}_0) + \mathbf{n}_{s,n},$$

where

$$E\{|\beta_{s,n}|^2\} = P,$$

the same for all subarrays,

$$E\{\mathbf{n}_{s,n}\mathbf{n}_{s,n}^H\} = \mathbf{I}_{M_s},$$

and

$$\mathbf{v}_s^H(\mathbf{u})\mathbf{v}_s(\mathbf{u}) = M_s,$$

for all $\mathbf{u}$, where M_s is the number of elements in the sth subarray. Assuming that the conditional probability density $p(\mathbf{z}|\mathbf{u})$ of the measurements is zero mean Gaussian, the logarithm of the likelihood function for the measurements taken from the sth subarray is [see Equations (8) and (9)]

$$\begin{aligned}\ln p(\mathbf{z}_{s,1}, \mathbf{z}_{s,2}, \ldots, \mathbf{z}_{s,N_s}|\mathbf{u}) &= -N_s[M_s \ln \pi + \ln(1 + M_s P)] \\ &\quad - \sum_{n=1}^{N_s} \mathbf{z}_{s,n}^H \left(\mathbf{I} - \frac{P}{1 + M_s P}\mathbf{v}_s(\mathbf{u})\mathbf{v}_s^H(\mathbf{u})\right)\mathbf{z}_{s,n}.\end{aligned}$$

The *maximum likelihood* estimate of $\mathbf{u}$ is that vector $\hat{\mathbf{u}}$, which maximizes the above expression. The only term that depends on $\mathbf{u}$ is

$$r_s(\mathbf{u}) = \frac{P}{1 + M_s P} \sum_{n=1}^{N_s} |\mathbf{z}_{s,n}^H \mathbf{v}_s(\mathbf{u})|^2, \tag{E.1}$$

the expected value of which is

$$E\{r_s(\mathbf{u})\} = \frac{P}{1+M_sP}\mathbf{v}_s^H(\mathbf{u})\sum_{n=1}^{N_s}E\{\mathbf{z}_{s,n}\mathbf{z}_{s,n}^H\}\mathbf{v}_s(\mathbf{u})$$

$$= \frac{P}{1+M_sP}(N_sP|\mathbf{v}_s^H(\mathbf{u})\mathbf{v}_s(\mathbf{u}_0)|^2 + M_s). \tag{E.2}$$

The maximum value of Equation (E.2) occurs at $\mathbf{u} = \mathbf{u}_0$. The *array power pattern* $g_s(\mathbf{u}, \mathbf{u}_0)$ of the sth subarray is defined to be the dot product in Equation (E.2) normalized so that its maximum value is unity, i.e.,

$$g_s(\mathbf{u}, \mathbf{u}_0) = |\frac{1}{M_s}\mathbf{v}_s^H(\mathbf{u})\mathbf{v}_s(\mathbf{u}_0)|^2.$$

The mean value of the log likelihood function for the entire set of switched measurements is just the sum of those for the individual subarrays, each of which is maximum when $\mathbf{u} = \mathbf{u}_0$. The quantity analogous to Equation (E.1) for switched subarrays is therefore

$$r(\mathbf{u}) = \sum_{s=1}^{S} r_s(\mathbf{u}), \tag{E.3}$$

with expected value

$$E\{r(\mathbf{u})\} = \sum_{s=1}^{S}\frac{M_sP}{1+M_sP}(N_sM_sPg_s(\mathbf{u}, \mathbf{u}_0) + 1).$$

The *switched subarray power pattern* is therefore defined to be a weighted sum of individual subarray patterns, the weights depending on the strength of the signal as well as the number of elements and looks contributed by each subarray.

$$g(\mathbf{u}, \mathbf{u}_0) = \sum_{s=1}^{S}\lambda_s g_s(\mathbf{u}, \mathbf{u}_0),$$

where

$$\lambda_s = \frac{N_sM_s^2}{1+M_sP}/\sum_{s=1}^{S}\frac{N_sM_s^2}{1+M_sP}.$$

For simplicity, attention will be confined here to subarrays of equal size ($M_s = M$), so that

$$\lambda_s = N_s / \sum_{s=1}^{S} N_s,$$

i.e., the weight of each subarray is just the fraction of the total number of looks it contributes. The sidelobe locations of the switched array pattern are used as test points when evaluating the WW bound on switched subarray DF error.

The estimate $\hat{\mathbf{u}}(\mathbf{z})$ that maximizes Equation (E.3) is the maximum likelihood estimate of direction for a switched subarray DF system. It is this estimator that was used in both the unswitched and switched simulations described in this report.

APPENDIX F
DERIVATION OF THE INFORMATION MATRIX FOR DF WITH PLANAR ARRAYS

This appendix derives an expression for the Fisher information matrix $\mathbf{F}$ for the array DF problem in the general case where each source is characterized by D real directional parameters. The directional parameters are treated as constant, unknown parameters. Although $D \leq 2$ for most DF problems, it is no harder to handle the general case, and the results may be useful in other applications.

Assume there are S signals (indexed by k,l) each having D parameters (indexed by m,n). The parameter vector for signal k is $\mathbf{u}_k^T = (u_{1k}, u_{2k}, \ldots, u_{Dk})$. The vector $\mathbf{u}$ is the concatenation of these vectors. The complex S-vector of signal amplitudes is assumed to be a zero mean Gaussian random vector with known covariance matrix $\mathbf{P}$.

The Fisher information matrix consists of S^2 blocks, each of dimension $D \times D$. The block corresponding to $\mathbf{u}_k, \mathbf{u}_l$ is (see Van Trees [1])

$$(\mathbf{F})_{\mathbf{u}_k\mathbf{u}_l^T} = E_{\mathbf{z}|\mathbf{u}}\{(\frac{\partial \ln p(\mathbf{z}|\mathbf{u})}{\partial \mathbf{u}_k})^T(\frac{\partial \ln p(\mathbf{z}|\mathbf{u})}{\partial \mathbf{u}_l})\}. \tag{F.1}$$

For brevity, denote the mnth element of $\mathbf{F}_{\mathbf{u}_k\mathbf{u}_l^T}$ by f_{mn}. Then (see DeLong [12, Appendix B])

$$f_{mn} = N \text{ Tr } (\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial u_{mk}}\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial u_{nl}}), \tag{F.2}$$

where

$$\mathbf{C}(\mathbf{u}) = \mathbf{I} + \mathbf{V}(\mathbf{u})\mathbf{P}\mathbf{V}^H(\mathbf{u})$$

is the multiple signal extension of Equation (7), in which

$$\mathbf{V}(\mathbf{u}) = (\mathbf{v}(\mathbf{u}_1), \mathbf{v}(\mathbf{u}_2), \ldots, \mathbf{v}(\mathbf{u}_S))$$

is a matrix, the columns of which are the array response vectors of the signals, and $\mathbf{P}$ is the covariance of the Gaussian signal amplitudes. The necessary derivative is

$$\frac{\partial \mathbf{C}}{\partial u_{mk}} = \frac{\partial \mathbf{V}(\mathbf{u})}{\partial u_{mk}}\mathbf{P}\mathbf{V}^H(\mathbf{u}) + \mathbf{V}(\mathbf{u})\mathbf{P}\frac{\partial \mathbf{V}^H(\mathbf{u})}{\partial u_{mk}}. \tag{F.3}$$

Only the kth column of $\mathbf{V}(\mathbf{u})$ depends on u_{mk}, so that

$$\frac{\partial \mathbf{V}(\mathbf{u})}{\partial u_{mk}} = (\mathbf{0}, \ldots, \mathbf{0}, \frac{\partial \mathbf{v}(\mathbf{u}_k)}{\partial u_{mk}}, \mathbf{0}, \ldots, \mathbf{0}) = \dot{\mathbf{V}}_k \mathbf{e}_m \mathbf{e}_k^T, \tag{F.4}$$

where

$$\dot{\mathbf{V}}_k \triangleq \frac{\partial \mathbf{v}(\mathbf{u}_k)}{\partial \mathbf{u}_k} \triangleq (\frac{\partial \mathbf{v}(\mathbf{u}_k)}{\partial u_{1k}}, \frac{\partial \mathbf{v}(\mathbf{u}_k)}{\partial u_{2k}}, \ldots, \frac{\partial \mathbf{v}(\mathbf{u}_k)}{\partial u_{Dk}}),$$

and $\mathbf{e}_m$ and $\mathbf{e}_k$ are unit vectors of dimension D and S, respectively, with unit element in row m and k, respectively.

Evaluation of Equation (F.2) using Equations (F.3) and (F.4) gives

$$\begin{aligned} f_{mn} &= N \text{ Tr } \{\mathbf{C}^{-1}(\dot{\mathbf{V}}_k \mathbf{e}_m \mathbf{e}_k^T \mathbf{P} \mathbf{V}^H + \mathbf{V}\mathbf{P}\mathbf{e}_k \mathbf{e}_m^T \dot{\mathbf{V}}_k^H)\mathbf{C}^{-1}(\dot{\mathbf{V}}_l \mathbf{e}_n \mathbf{e}_l^T \mathbf{P} \mathbf{V}^H + \mathbf{V}\mathbf{P}\mathbf{e}_l \mathbf{e}_n^T \dot{\mathbf{V}}_l^H)\} \\ &= N \text{ Tr } \{(\mathbf{e}_k^T \mathbf{P} \mathbf{V}^H \mathbf{C}^{-1} \dot{\mathbf{V}}_l \mathbf{e}_n)(\mathbf{e}_l^T \mathbf{P} \mathbf{V}^H \mathbf{C}^{-1} \dot{\mathbf{V}}_k \mathbf{e}_m) + \text{complex conjugate} + \\ &\qquad (\mathbf{e}_m^T \dot{\mathbf{V}}_k^H \mathbf{C}^{-1} \dot{\mathbf{V}}_l \mathbf{e}_n)(\mathbf{e}_l^T \mathbf{P} \mathbf{V}^H \mathbf{C}^{-1} \mathbf{V}\mathbf{P}\mathbf{e}_k) + \text{complex conjugate } \} \\ &= 2N \, \Re\{(\mathbf{P}\dot{\mathbf{W}}_l)_{kn}(\mathbf{P}\dot{\mathbf{W}}_k)_{lm} + (\ddot{\mathbf{W}}_{kl})_{mn}(\mathbf{P}\mathbf{W}\mathbf{P})_{lk},\} \end{aligned} \tag{F.5}$$

where

$$\mathbf{W} \triangleq \mathbf{V}^H \mathbf{C}^{-1} \mathbf{V}, \tag{F.6}$$

$$\dot{\mathbf{W}}_k \triangleq \mathbf{V}^H \mathbf{C}^{-1} \dot{\mathbf{V}}_k, \tag{F.7}$$

$$\ddot{\mathbf{W}}_{kl} \triangleq \dot{\mathbf{V}}_k^H \mathbf{C}^{-1} \dot{\mathbf{V}}_l. \tag{F.8}$$

The dot notation for the $\mathbf{W}$ is for mnemonic purposes only; it does not indicate differentiation. Let $(\mathbf{A}_{kl})_{mn}$ and $(\mathbf{B}_{kl})_{mn}$ denote the first and second terms inside the brackets in Equation (F.5).

Define

$$\dot{\mathbf{W}} \triangleq (\dot{\mathbf{W}}_1, \dot{\mathbf{W}}_2, \ldots, \dot{\mathbf{W}}_S)$$

and

$$\ddot{\mathbf{W}} \triangleq \begin{pmatrix} \ddot{\mathbf{W}}_{11} & \ddot{\mathbf{W}}_{12} & \cdots & \ddot{\mathbf{W}}_{1S} \\ \ddot{\mathbf{W}}_{21} & \ddot{\mathbf{W}}_{22} & \cdots & \ddot{\mathbf{W}}_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ \ddot{\mathbf{W}}_{S1} & \ddot{\mathbf{W}}_{S2} & \cdots & \ddot{\mathbf{W}}_{SS} \end{pmatrix}.$$

The second term $(\mathbf{B}_{kl})_{mn}$ is recognized as the product of the mnth element of $\ddot{\mathbf{W}}_{kl}$ and the scalar $c_{kl} = ((\mathbf{PWP})^T)_{kl}$, so the contribution of this term to the complete Fisher information matrix $\mathbf{F}$ is

$$\mathbf{B} = \begin{pmatrix} \ddot{\mathbf{W}}_{11}c_{11} & \ddot{\mathbf{W}}_{12}c_{12} & \cdots & \ddot{\mathbf{W}}_{1S}c_{1S} \\ \ddot{\mathbf{W}}_{21}c_{21} & \ddot{\mathbf{W}}_{22}c_{22} & \cdots & \ddot{\mathbf{W}}_{2S}c_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ \ddot{\mathbf{W}}_{S1}c_{S1} & \ddot{\mathbf{W}}_{S2}c_{S2} & \cdots & \ddot{\mathbf{W}}_{SS}c_{SS} \end{pmatrix}$$

$$= \ddot{\mathbf{W}} \square \begin{pmatrix} c_{11}\mathbf{1}_D\mathbf{1}_D^T & c_{12}\mathbf{1}_D\mathbf{1}_D^T & \cdots & c_{1S}\mathbf{1}_D\mathbf{1}_D^T \\ c_{21}\mathbf{1}_D\mathbf{1}_D^T & c_{22}\mathbf{1}_D\mathbf{1}_D^T & \cdots & c_{2S}\mathbf{1}_D\mathbf{1}_D^T \\ \vdots & \vdots & \ddots & \vdots \\ c_{S1}\mathbf{1}_D\mathbf{1}_D^T & c_{S2}\mathbf{1}_D\mathbf{1}_D^T & \cdots & c_{SS}\mathbf{1}_D\mathbf{1}_D^T \end{pmatrix}$$

$$= \ddot{\mathbf{W}} \square ((\mathbf{PWP})^T \otimes \mathbf{1}_D\mathbf{1}_D^T). \tag{F.9}$$

The symbol $\square$ denotes the Hadamard product;[8] the symbol $\otimes$ denotes the Kronecker product.[9]

The first term can be written as

$$(\mathbf{A}_{kl})_{mn} = (\mathbf{P}\dot{\mathbf{W}}_k)^T_{ml}(\mathbf{P}\dot{\mathbf{W}}_l)_{kn} = \mathbf{e}_m^T(\mathbf{P}\dot{\mathbf{W}}_k)^T\mathbf{e}_l\mathbf{e}_k^T(\mathbf{P}\dot{\mathbf{W}}_l)\mathbf{e}_n,$$

---

[8] $(\mathbf{A}\square\mathbf{B})_{mn} \triangleq a_{mn}b_{mn}$. The Hadamard product is defined only when A and B have the same dimensions.

[9] $\mathbf{A}\otimes\mathbf{B} \triangleq \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots \\ \cdots & \cdots & \ddots \end{pmatrix}$. The Kronecker product is defined for any two matrices A and B.

showing that the klth block of $\mathbf{A}$ is

$$\mathbf{A}_{kl} = (\mathbf{P}\dot{\mathbf{W}}_k)^T \mathbf{e}_l \mathbf{e}_k^T \mathbf{P}\dot{\mathbf{W}}_l, \tag{F.10}$$

which is a dyad involving two D-dimensional vectors. Any two such vectors $\mathbf{x}$ and $\mathbf{y}$ satisfy the identity

$$\mathbf{x}\mathbf{y}^T \equiv (\mathbf{x}\mathbf{1}_D^T) \square (\mathbf{1}_D \mathbf{y}^T).$$

Thus Equation (F.10) becomes

$$\mathbf{A}_{kl} = ((\mathbf{P}\dot{\mathbf{W}}_k)^T \mathbf{e}_l \mathbf{1}_D^T) \square (\mathbf{1}_D \mathbf{e}_k^T \mathbf{P}\dot{\mathbf{W}}_l).$$

Define

$$\mathbf{Q}_{kl} = (\mathbf{P}\dot{\mathbf{W}}_k)^T \mathbf{e}_l \mathbf{1}_D^T$$

and observe that

$$\mathbf{A}_{kl} = \mathbf{Q}_{kl} \square (\mathbf{Q}_{lk})^T,$$

so that the complete $\mathbf{A}$ matrix is of the form

$$\mathbf{A} = \mathbf{Q} \square \mathbf{Q}^T.$$

The full $\mathbf{Q}$ matrix has the form

$$\mathbf{Q} = \begin{pmatrix} (\mathbf{P}\dot{\mathbf{W}}_1)^T \mathbf{e}_1 \mathbf{1}_D^T & (\mathbf{P}\dot{\mathbf{W}}_1)^T \mathbf{e}_2 \mathbf{1}_D^T & \cdots & (\mathbf{P}\dot{\mathbf{W}}_1)^T \mathbf{e}_S \mathbf{1}_D^T \\ (\mathbf{P}\dot{\mathbf{W}}_2)^T \mathbf{e}_1 \mathbf{1}_D^T & (\mathbf{P}\dot{\mathbf{W}}_2)^T \mathbf{e}_2 \mathbf{1}_D^T & \cdots & (\mathbf{P}\dot{\mathbf{W}}_2)^T \mathbf{e}_S \mathbf{1}_D^T \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbf{P}\dot{\mathbf{W}}_S)^T \mathbf{e}_1 \mathbf{1}_D^T & (\mathbf{P}\dot{\mathbf{W}}_S)^T \mathbf{e}_2 \mathbf{1}_D^T & \cdots & (\mathbf{P}\dot{\mathbf{W}}_S)^T \mathbf{e}_S \mathbf{1}_D^T \end{pmatrix}$$

$$= \left((\mathbf{P}\dot{\mathbf{W}})^T \mathbf{e}_1 \mathbf{1}_D^T, (\mathbf{P}\dot{\mathbf{W}})^T \mathbf{e}_2 \mathbf{1}_D^T, \ldots, (\mathbf{P}\dot{\mathbf{W}})^T \mathbf{e}_S \mathbf{1}_D^T \right)$$

$$= (\mathbf{P}\dot{\mathbf{W}})^T(\mathbf{e}_1\mathbf{1}_D^T, \mathbf{e}_2\mathbf{1}_D^T, \ldots, \mathbf{e}_S\mathbf{1}_D^T).$$

The matrix on the right above can be written more compactly as $\mathbf{I}_S \otimes \mathbf{1}_D^T$. The complete $\mathbf{A}$ matrix can then be written as

$$\mathbf{A} = ((\mathbf{P}\dot{\mathbf{W}})^T(\mathbf{I}_S \otimes \mathbf{1}_D^T))\square((\mathbf{I}_S \otimes \mathbf{1}_D)\mathbf{P}\dot{\mathbf{W}}). \tag{F.11}$$

Finally, Equations (F.11) and (F.9) in combination give the complete Fisher information matrix for the directional parameters,

$$(\mathbf{F})_{\mathbf{u}\mathbf{u}^T} = 2N\,\Re\{\ddot{\mathbf{W}}\square((\mathbf{P}\mathbf{W}\mathbf{P})^T\otimes\mathbf{1}_D\mathbf{1}_D^T)+((\mathbf{P}\dot{\mathbf{W}})^T(\mathbf{I}_S\otimes\mathbf{1}_D^T))\square((\mathbf{I}_S\otimes\mathbf{1}_D)\mathbf{P}\dot{\mathbf{W}})\}. \tag{F.12}$$

If the array elements are omnidirectional and only one signal is present, Equation (F.12) becomes quite simple. $\mathbf{P}$ and

$$\mathbf{W} = \mathbf{v}^H(\mathbf{I} - \frac{P}{MP+1}\mathbf{v}\mathbf{v}^H)\mathbf{v} = \frac{M}{MP+1}$$

become scalars, and because

$$\dot{\mathbf{V}} \triangleq (\frac{\partial \mathbf{v}(\mathbf{u})}{\partial u_x}, \frac{\partial \mathbf{v}(\mathbf{u})}{\partial u_y}) = jk[\mathbf{v}(\mathbf{u})]\mathbf{D},$$

where $\mathbf{D}$ is given by Equation (22), Equations (F.7) and (F.8) become

$$\dot{\mathbf{W}} = jk\mathbf{v}^H(\mathbf{I} - \frac{P}{MP+1}\mathbf{v}\mathbf{v}^H)[\mathbf{v}]\mathbf{D} = \frac{jk}{MP+1}\mathbf{1}^T\mathbf{D},$$

$$\ddot{\mathbf{W}} = k^2\mathbf{D}^T[\mathbf{v}]^H(\mathbf{I} - \frac{P}{MP+1}\mathbf{v}\mathbf{v}^H)[\mathbf{v}]\mathbf{D} = k^2\mathbf{D}^T(\mathbf{I} - \frac{P}{MP+1}\mathbf{1}\mathbf{1}^T)\mathbf{D}.$$

The first matrix inside the brackets in Equation (F.12) becomes

$$\frac{k^2MP^2}{MP+1}\mathbf{D}^T(\mathbf{I} - \frac{P}{MP+1}\mathbf{1}\mathbf{1}^T)\mathbf{D},$$

while the second becomes

$$-\frac{k^2P^2}{(MP+1)^2}\mathbf{D}^T\mathbf{1}\mathbf{1}^T\mathbf{D}.$$

The result is then just

$$\mathbf{F} = \frac{2Nk^2MP^2}{MP+1}\mathbf{D}^T(\mathbf{I} - \frac{1}{M}\mathbf{1}\mathbf{1}^T)\mathbf{D}.$$

Note that this result is not a function of $\mathbf{u}$.

REFERENCES

1. H.L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, New York, NY: John Wiley and Sons (1968).
2. E. Weinstein and A.J. Weiss, "A General Class of Lower Bounds in Parameter Estimation," *IEEE Transactions on Information Theory* **34**, 338-342 (1988).
3. Ibid, "Lower Bounds on the Mean Square Estimation Error," in *Proc. of the IEEE* **73**, 1433-1434 (1985).
4. A.J. Weiss and E. Weinstein, "A Lower Bound on the Mean-Square Error in Random Parameter Estimation," *IEEE Transactions on Information Theory* **31**, 680-682 (1985).
5. D.H. Johnson, "The Application of Spectral Estimation Methods to Bearing Estimation Problems," in *Proc. of the IEEE* **70**, 1018-1028 (1982).
6. H. Cramer, *Mathematical Methods of Statistics*, Princeton, NJ: Princeton University Press (1951).
7. D.F. DeLong, "Multiple Signal Direction Finding With Thinned Linear Arrays," MIT Lincoln Laboratory, Lexington, Mass, Project Rep. TST-68 (13 April 1983).
8. F.M. White, "Performance of Bayes-Optimal Angle-of-Arrival Estimators," MIT Lincoln Laboratory, Lexington, Mass, Technical Rep. 654 (13 August 1984).
9. J.E. Evans, J.R. Johnson, and D.F. Sun, "Application of Advanced Signal Processing Techniques to Angle of Arrival Estimation in ATC Navigation and Surveillance Systems," MIT Lincoln Laboratory, Lexington, Mass, Technical Rep. 582 (23 June 1982).
10. R.J. McAulay and L.P. Seidman, "A Useful Form of the Barankin Lower Bound and Its Application to PPM Threshold Analysis," *IEEE Transactions on Information Theory* **15**, 273-279 (1969).
11. W. Rudin, *Real and Complex Analysis*, New York, NY: McGraw-Hill (1974).
12. D. F. DeLong, "A New Algorithm for Adaptive Radar Signal Processing," MIT Lincoln Laboratory, Lexington, MA, Technical Note 1979-72 (31 December 1979). AD-A082856

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	16 August 1993	Technical Report

4. TITLE AND SUBTITLE

Use of the Weiss-Weinstein Bound to Compare the Direction-Finding Performance of Sparse Arrays

5. FUNDING NUMBERS

C — F19628-90-C-0002
PR — 589

6. AUTHOR(S)

Darrol F. DeLong

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Lincoln Laboratory, MIT
P.O. Box 73
Lexington, MA 02173-9108

8. PERFORMING ORGANIZATION REPORT NUMBER

TR-982

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Department of Defense
Ft. George G. Meade, MD 20755

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

ESC-TR-93-244

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

Weiss and Weinstein have recently derived a new lower bound on random parameter estimation error. This report makes use of this bound to develop a criterion for comparing two sparse array configurations. The arrays are first adjusted in size so that their direction-finding (DF) accuracies at high element signal-to-noise ratio (ESNR) are the same when averaged over all possible directions of arrival in the hemisphere. Then their threshold ESNRs are calculated using the Weiss-Weinstein bound. The array having the lower threshold ESNR is judged the better array. As an example of the application of the bound, three 7-element arrays, a circle, a spiral, and an L-shaped array, are compared.

The bound is also extended to the case of switched subarrays, where data are collected from several subarrays sequentially in time and processed as a unit. The purpose is to achieve DF accuracy approaching that of the full array while reducing the number of signal processing channels. An example, using five receivers on the 7-element spiral array, shows that this is indeed possible.

14. SUBJECT TERMS		15. NUMBER OF PAGES
Cramer-Rao factor	element signal-to-noise ratio	66
Weiss-Weinstein bound	switched arrays	16. PRICE CODE
direction finding		

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Same As Report